REMARKS

Claims 1-6 and 8-27 are currently pending in the subject application and are presently under consideration. Claims 22-27 have been added and claims 1, 4, 9, 10, 12, 14, and 16-21 have been amended as shown on pp. 2-9 of the Reply. Claims 7, 11, 13, and 15 have been canceled without prejudice or disclaimer. Further, Applicants' representative notes with appreciation that the rejection of claims 1-19 under 35 U.S.C. §101 has been overcome, as indicated in the Advisory Action dated September 14, 2007.

Favorable reconsideration of the subject patent application is respectfully requested in view of the comments and amendments herein.

I. New Claims 22-27

In the interest of expedited prosecution, the following provides locations in the specification at which support can be found for the limitations of new claims 27-27.

As set forth above, new independent claim 27 recites:

A method of modeling speech dynamics for a speech processing application, comprising:

constructing a speech model, the speech model is based on a hidden dynamic model in the form of a segmental switching state space model for speech applications (p. 5, ll. 28-29)*, the constructing a speech model comprising: initializing a first set of model parameters that describes unobserved vocal tract resonance frequencies* (p. 6, ll. 7-8)*; initializing a second set of model parameters that describes a mapping relationship between the unobserved vocal tract resonance frequencies and observed acoustic data* (p. 5, ll. 24-26)*; creating a state equation based on the first set of model parameters to express the unobserved vocal tract resonance frequencies as a set of states respectively corresponding to phones in an unobserved phonetic transcript, the state equation is a linear dynamic equation that describes transitions between states in the set of states in terms of a phone-dependent system matrix and a target vector and includes a first Gaussian noise parameter* (p. 6, ll. 8-10; p. 6, ll. 14-15)*; creating an observation equation that utilizes the first set of model parameters and the second set of model parameters to represent a phone-dependent mapping between the unobserved vocal tract resonance frequencies and the observed acoustic data, the mapping selected from the*

group consisting of a linear mapping and a piecewise linear mapping within respective phones, the observation equation includes a second Gaussian noise parameter (p. 6, ll. 10-15)*; estimating soft phone boundaries for phones in the unobserved phonetic transcript under an expectation-maximization (EM) framework* (p. 7, ll. 26-27)*; and constructing a series of time-varying transition matrices based on the phonetic transcript to constrain the set of states to respective time durations corresponding to the estimated soft phone boundaries for phones in the phonetic transcript, thereby forcing the states to be consistent in time with the phonetic transcript* (p. 14, ll. 9-14)*;*

calculating an estimated multimodal posterior distribution based on the constructed speech model, the first set of model parameters, and the second set of model parameters (p. 7, ll. 27-29)*; and*

modifying one or more model parameters to minimize a Kullback-Leibler distance from the estimated multimodal posterior distribution to an exact posterior distribution (p. 7, ll. 21-23)*, the modifying is based on an EM framework having an expectation step of model inference and a maximization step of model learning* (p. 7, ll. 9-10)*, the model learning is based on a variational learning technique that employs calculus of variation* (p. 7, ll. 13-17; p. 8, l. 18-19).

With regard to new dependent claims 22-26, said claims find support in the specification at similar locations to new independent claim 27.

II. <u>Rejection of Claims 1-5, 7-13, and 19-21 Under 35 U.S.C. §103(a)</u>

Claims 11-5, 7-13, and 19-21 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Hogden (US 6,052,662) in view of Ghahramani *et al.*, "Variational Learning for Switching State-Space Models" (Neural Computation 2000). Withdrawal of this rejection is requested for at least the following reasons. The cited references, either alone or in combination, do not disclose or suggest all features recited in the subject claims as amended. "To reject claims in an application under §103 . . . the prior art reference (or references when combined) must teach or suggest all the claim limitations." *See* MPEP §706.02(j); *see In re Vaeck,* 947 F.2d 488, 20 USPQ2d 1438 (Fed. Cir. 1991).

Amended independent claim 1 (and its corresponding dependent claims) recites: *A system that facilitates speech recognition by modeling speech dynamics, comprising: an*

input component that receives acoustic data; and a model component that employs the acoustic data to characterize speech, the model component comprising model parameters that form a mapping relationship from unobserved speech dynamics to observed speech acoustics, the model parameters are employed to decode an unobserved phone sequence of speech based, at least in part, upon a variational learning technique; ***wherein the model component is based, at least in part, upon a hidden dynamic model in the form of a segmental switching state space model, the segmental switching state space model comprises respective states having respective durations in time corresponding to soft boundaries of respective phones in the unobserved phone sequence.*** The subject amendments are supported by the specification. For example, the specification discloses that segmental constraints can be applied to a speech model in order to force states used by the model to be consistent in time with a phonetic transcript. (*See* p. 14, ll. 9-14). Further, the specification discloses that estimated soft phone assignments can be utilized by the model to facilitate recovery of a phone sequence. (*See* p. 7, ll. 26-29).

Hogden relates to a speech processing methodology called Maximum Likelihood Continuity Mapping (Malcom), which models acoustic speech data as a continuous pseudo-articular path. (*See, e.g.,* col. 5, ll. 10-13). Malcom determines a pseudo-articular path for a given set of acoustic speech data by finding the pseudo-articular path that would be most likely to produce the acoustic speech data. (*See, e.g.,* col. 8, ll. 31-37). However, as conceded by the Examiner on page 5 of the Office Action, Hogden does not disclose the use of a hidden dynamic model in the form of a segmental switching state space model. To overcome this deficiency of Hogden, the Examiner cites Ghahramani *et al.* Said reference relates to the creation and use of segmental switching state space models for applications in fields such as econometrics and signal processing. (*See, e.g.,* p. 1, para. 5). In addition, Ghahramani *et al.* describes two experiments performed using segmental switching state space models. The first of these experiments was performed on artificial test data generated by two state-space models. (*See* Section 5.1; p. 12, para. 6). The second of these experiments was performed on respiration force data obtained from a person with sleep apnea. (*See* Section 5.2; p. 13, para. 3). However, independent claim 1 recites that ***the model component is based, at least in part, upon a hidden dynamic model in the form of a segmental switching state space model, the segmental***

switching state space model comprises respective states having respective durations in time corresponding to soft boundaries of respective phones in the unobserved phone sequence. The cited references do not disclose or suggest such features.

While Ghahramani et al. discloses employing a model comprising discrete states for signal processing, said reference is silent as to employing a model comprising discrete states having respective durations in time corresponding to soft phone boundaries, as recited by independent claim 1. In particular, Ghahramani *et al.* discloses two examples of segmental constraints that can be used to generate discrete states for a data model. In the first example, a segmental switching state space model receives an input signal created from two state space models and divides the input signal into the segments produced by the first model and the segments produced by the second model. (*See* Section 5.1; p. 12, para. 6). In the second example, a segmental switching state space model divides an input signal corresponding to respiration force into segments corresponding to periods of rhythmic breathing and segments corresponding to periods of apnea. (*See* Section 5.2; p. 13, para. 3). However, neither of these examples is sufficient to suggest segmentation of model states based on phone boundaries.

In both of the examples disclosed in Ghahramani *et al.*, segmentation was based on input data with two well-defined states. On the other hand, a segmental switching state space model for speech data, such as the model recited by independent claim 1, requires segmentation based on phone boundaries. (*See* p. 7, ll. 26-27). The number of phones possible in human speech clearly far exceeds the two states on which the segmentation in the examples described in Ghahramani *et al.* was based. Further, segmentation based on phone boundaries must account for much more subtle differences in an input data stream than the differences between states presented by the well-defined states utilized in the examples given in Ghahramani *et al.* The subtle differences between states based on phones in a phone sequence, as recited by independent claim 1, demonstrate that applying a segmental switching state space model to a speech application would involve adaptation of a data model beyond the teachings and/or suggestions of Hogden and Ghahramani *et al.* Thus, the cited references, either alone or in combination, do not disclose or suggest all limitations of independent claim 1.

Independent claims 12 and 19-21 have been amended to recite similar features, namely *a model based, at least in part, upon a hidden dynamic model in the form of a segmental switching state space model, the segmental switching state space model comprises respective states having respective durations in time corresponding to soft boundaries of respective phones in the unobserved phone sequence.* Accordingly, the cited references, either alone or in combination, do not disclose or suggest all limitations of independent claims 12 and 19-21 for the reasons stated above regarding independent claim 1. In view of the foregoing, applicants' representative respectfully requests that this rejection be withdrawn.

III. Rejection of Claims 6 and 14-18 Under 35 U.S.C. §103(a)

Claims 6 and 14-18 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Hogden in view of Ghahramani *et al.* and further in view of McDonough (US 5,652,748). Withdrawal of this rejection is requested for at least the following reasons.

With regard to claim 6, Applicants' representative notes that independent claim 1, from which this claim depends, has been amended to recite features not disclosed or suggested by Hogden or Ghahramani *et al.* Further, McDonough does not cure the deficiencies of said references with regard to independent claim 1. Thus, the cited references, either alone or in combination, do not disclose or suggest all limitations of claim 6.

In addition, independent claims 14 and 17 (and their corresponding dependent claims) have been amended in a similar manner to independent claim 1 to include a segmental switching state space model comprising states having respective durations in time corresponding to soft boundaries of phones in a recovered phone sequence, which is not disclosed or suggested by Hogden or Ghahramani *et al.* Further, McDonough does not cure the deficiencies of said references with regard to independent claims 14 and 17. Accordingly, the cited references, either alone or in combination, do not teach or suggest all limitations of claims 14-18. In view of the foregoing, Applicants' representative respectfully requests that this rejection be withdrawn.

CONCLUSION

The present application is believed to be in condition for allowance in view of the above comments and amendments. A prompt action to such end is earnestly solicited.

In the event any fees are due in connection with this document, the Commissioner is authorized to charge those fees to Deposit Account No. 50-1063 [MSFTP435US].

Should the Examiner believe a telephone interview would be helpful to expedite favorable prosecution, the Examiner is invited to contact applicants' undersigned representative at the telephone number below.

Respectfully submitted,
Amin, Turocy & Calvin, LLP

/Himanshu S. Amin/
Himanshu S. Amin
Reg. No. 40,894

Amin, Turocy & Calvin, LLP
24TH Floor, National City Center
1900 E. 9TH Street
Cleveland, Ohio 44114
Telephone (216) 696-8730
Facsimile (216) 696-8731